BMC Plant Biology

RESEARCH

Open Access

Comparative and phylogenetic analyses based on the complete chloroplast genomes of *Limonium* (Plumbaginaceae) species

Fangfang Jiao[1], Gulbar Yisilam[1,2,3], Aiqin Zhang[1*], Xiaohong Yang[1], Jing Zhang[1] and Xinmin Tian[2,3*]

Abstract

Background *Limonium* Mill., a genus belonging to the family Plumbaginaceae, is mainly distributed along the Mediterranean coast of Eurasia. It is known for its rich species diversity and remarkable morphological variation, encompassing both herbaceous and semi-shrub forms. Some species of *Limonium* have medicinal, economic, and ornamental value in China. However, their genomic and phylogenetic relationships remain unclear. In this study, we sequenced and assembled complete chloroplast genomes of eight *Limonium* species to address this issue. We conducted comparative and phylogenetic analyses to investigate the genetic structural characteristics and clarify the taxonomic status of 12 *Limonium* species.

Results The newly obtained circular chloroplast genomes of the eight *Limonium* species had a quadripartite structure and ranged from 154,545 to 154,960 bp in length. All species contained 128 genes, including 83 protein-coding genes, 37 tRNA genes, and 8 rRNA genes. Despite minor variations in the inverted repeat (IR) boundary regions, the overall genome structure and gene content remained relatively conserved. The *rpl16* gene lost its intron and the *rpl23*gene underwent pseudogenization. We identified 55–73 SSRs repeat sequences and three nucleotide diversity regions that will be used for population genetics and molecular phylogenetic studies of the *Limonium* genus. The robust phylogenetic tree suggests that species belonging to these sections in Xinjiang, China, have accumulated substantial genetic divergence over long-term evolution, resulting in stable phylogenetic relationships.

Conclusion The complete chloroplast genomes of *Limonium* increase our understanding of the genetic diversity and evolution of *Limonium* species and further aid in the exploration and utilization of *Limonium* plants.

Keywords *Limonium*, Medicinal plant, Chloroplast genomes, Comparative analysis, Phylogenomics

*Correspondence:
Aiqin Zhang
zhangaq@xju.edu.cn
Xinmin Tian
tianxm333333@foxmail.com
[1]Xinjiang Key Laboratory of Biological Resources and Genetic Engineering, College of Life Science and Technology, Xinjiang University, Urumqi 830017, China
[2]Key Laboratory of Ecology of Rare and Endangered Species and Environmental Protection (Ministry of Education) & Guangxi Key Laboratory of Landscape Resources Conservation and Sustainable Utilization in Lijiang River Basin, Guangxi Normal University, Guilin 541006, China
[3]University Engineering Research Center of Bioinformation and Genetic Improvement of Specialty Crops, Guilin 541006, China

Background

Limonium L., a member of the Plumbaginaceae family, is among the largest genera worldwide and comprises nearly 600 species [1, 2]. Approximately 22 species of *Limonium* have been identified in China, primarily in Xinjiang [3]. These plants predominantly thrive in soils that are both saline and rich in metals across various regions, including the mainland and coastal areas [4]. They are recognized as pioneering species in the saline soils of China, Russia, and Mongolia. A case in point is *L. bicolor*, which serves as a model recretohalophyte [5]. It thrives and can complete its life cycle in highly saline soils because of salt glands on its epidermis [6] and a thick cuticle, which reduces water evaporation [7]. Its salt and drought tolerance enables it to maintain ecological balance in ecosystems affected by environmental stress, thereby contributing to the sustainable development of regional ecology. Beyond their ecological significance, *Limonium* species hold considerable medicinal and economic value. It is a source of bioactive compounds used extensively in traditional and contemporary medicine [8]. They exhibit potential anti-metastatic properties, counteract melanin activity, inhibit cell proliferation pathways, and possess antioxidant and immunomodulatory effects [9–12]. They are considered potential sources of anti-HIV-1 drugs [13]. Furthermore, *Limonium*, with its persistent calyx, is widely used as a dried flower and is also widely used in flower bouquets because of its colorful colors [14, 15].

The taxonomy and phylogenetic relationships within *Limonium* are notoriously complex. Although many new species of *Limonium* have been discovered in the Mediterranean region over the past half-century [1, 2], the diversity in reproductive mechanisms, including sexual reproduction and apomixis, and with hybridization and the rise of polyploidy, have complicated the clear delimitation of many *Limonium* species [4, 16–19]. Malekmohammadi et al. [20] addressed the phylogenetic relationships within *Limonium*, focusing on 76 species, using the nuclear internal transcribed spacer (ITS) region and several chloroplast DNA datasets. Their study identified two well-supported clades corresponding to the subgenera *Limonium* and *Pteroclados* (Boiss.) Pignatti, corroborating earlier findings by Lledó et al. [21, 22]. Koutroumpa et al. [1] conducted a comprehensive phylogenetic analysis of the Plumbaginaceae family, including 23 genera, with a particular focus on *Limonium*. They affirmed the majority of prior molecular phylogenetic findings and prompted the introduction of novel sections, as well as revisions to existing classifications. To date, the most comprehensive phylogenetic relationship analysis within *Limonium* has encompassed 214 species worldwide; however, various subclades and infrageneric relationships exhibit weak support [23]. The scarcity of genomic resource information hinders a clear understanding of their evolutionary history. Therefore, delving deeper into *Limonium* using a more holistic genomic approach, including chloroplast genome analysis, is highly pertinent and could significantly enhance our understanding of this domain.

The chloroplast genome has emerged as a cornerstone for plant evolutionary and phylogenetic studies. It's generally conserved structure, coupled with variable regions, provides valuable insights into relationships among species [24]. Typically ranging in size from 120 to 160 kilobase pairs (kb), the chloroplast genome is independently inherited and offers a rich source of genetic information for phylogenetics and DNA barcode development [25–27]. Scholars have increasingly favored the use of chloroplast genomes to address issues related to species identification, phylogenetic relationships among different species, systematics, and evolution, and to identify polyploidy events of genera [28–31]. While chloroplast genomic sequences for a few *Limonium* species are available [32–35]. However, our understanding of the genetic content and structural diversity within this genus remains somewhat constrained. Moreover, studying the chloroplast genome is beneficial for developing chloroplast genomic resources in *Limonium*, elucidating the phylogenetic relationships among different species within the genus and identifying potential DNA markers that are crucial for future research on population genetics or the phylogeography of this genus.

In the present study, eight *Limonium* species were chosen to represent key taxonomic sections within the genus found in China, including species with ecological importance, medicinal value (e.g., *L. aureum* and *L. gmelinii*) and pharmacological studies remain essentially nonexistent (e.g., *L. chrysocomum* and *L. suffruticosum*). Moreover, some of these species have not been recorded elsewhere in China, except for sporadic occurrences in disparate habitats of northern and southern Xinjiang, China. The widespread species, endemic species to the region, and species with narrow distribution areas collectively encompass most of the morphological and ecological diversity of the genus in China. Therefore, we chose to elucidate the chloroplast genome characteristics of these species in order to understand that their genetic resources are crucial for guiding future conservation, evolution and potential utilization studies. We sequenced and assembled the complete chloroplast genomes of these species. Using bioinformatics, we conducted comparative genomic and phylogenetic analyses of the complete chloroplast genomes across 12 *Limonium* taxa. Our aims were to (1) analyze complete chloroplast genomic features in *Limonium*, (2) identify highly variable regions that could serve as molecular markers and potentially specific barcodes for *Limonium*, and (3) provide insights

into the phylogenetic relationships among *Limonium* species. The findings of this study are expected to bolster ongoing studies on species identification, population genetics, phylogenetics, and the evolution of complete chloroplast genomes within *Limonium*. Additionally, this study provides a theoretical foundation for conserving *Limonium* species.

Results

Sequencing, Chloroplast genome structure, and characteristics

Chloroplast genomes length ranged from 154,545 to 154,960 bp in the eight *Limonium* species. These chloroplast genomes were composed of a large single-copy region (LSC) between 84,427 and 84,651 bp, a small single-copy (SSC) region between 12,980 and 13,160 bp, and two inverted repeat (IR) regions spanning 28,535 to 28,623 bp (Fig. 1). A total of 128 genes were annotated, including 16 duplicates and 112 unique genes. The 112 unique genes comprised 78 protein-coding genes, 4 ribosomal RNA (rRNA) genes, and 30 transfer RNA (tRNA) genes. Among the 112 unique genes, 11 protein-coding genes and 5 tRNA genes had a single intron, whereas 2 genes contained two introns (Table S1). The intron of *rpl16* was lost, and *rpl23* was prematurely interrupted by internal stop codons. The overall guanine-cytosine (GC) content varied from 36.9% to 37.1%, with GC contents between 35.2% and 35.4% in LSC and from 31.2% to 33.6% in SSC. The GC content was relatively high in the IR region, ranging from 40.8% to 42.8%, surpassing that in the LSC and SSC regions (Table 1).

Comparative chloroplast genomic analysis

Comparative genomic analysis indicated a relatively high conservation among the 12 *Limonium* chloroplast genomic sequences (Fig. 2A). Most protein-coding regions were notably conserved, and rRNA genes displayed minimal divergence across species. Surprisingly, *L. tenellum* underwent massive variation from 104 kb to 116 kb.

The nucleotide diversity (Pi) values ranged from 0 to 0.04124, with the *trnQ-UUG-psbK* interval showing maximum diversity at 0.04321. The average Pi values observed in the IR regions were lower than those in the SSC and LSC regions (Fig. 2B), indicating that the IR regions were more conserved than the single-copy regions (SC). We identified three regions with high variability (Pi > 0.04) in the SC: *trnQ-UUG-psbK*, *rpl32-trnL-UAG*, and *rps16-trnQ-UUG* (Fig. 2B).

A comparison of the 12 chloroplast genomes revealed minor variations in the IR region boundaries (Fig. 2C). The IRb/LSC junction of *Limonium* species was identified by the presence of the *rpl2* gene, which varied slightly in size between 1439 bp and 1440 bp within the LSC region. The gene also extended into the IRb region by 10–120 bp. The IRb/SSC junction was defined by the *ndhF* gene, which extended into the IRb by 2,177 bp or 2,171 bp. In *L. tenellum*, this junction displayed the components *rrn4* (IR) and *trnR-ACG* (SSC), likely due to assembly errors. At the IRa/SSC, the *rps15* gene, which is fully located in the IRa region, experienced a slight reduction of 3 bp. The *trnH* genes were enlarged from the LSC region into the IRa region with an increase of 2–4 bp. *L. tenellum* had a single copy, with the *ycf1* gene spanning 5,298 bp, and the IR region was enlarged to include this gene, whereas all other species possessed two copies of the *ycf1* gene (Fig. 2C).

Repeat sequences and codon usage bias analysis

A significant variation in the quantity of simple sequence repeats (SSRs) was observed within the chloroplast genomes of 12 *Limonium* species in repetitive sequence analysis, with counts varying from 55 in *L. tenellum* to 73 in *L. suffruticosum*. Among these SSRs, 457 A/T bp accounted for the highest proportion (59.82%), whereas C and G nucleotides had the lowest proportion at 2.62%. Notably, SSRs showed diverse nucleotide variations, with di-, tri-, tetra-, penta-, and hexanucleotide repeats accounting for 18.46%, 7.75%, 10.34%, 0.78%, and 0.26% of the total repeats, respectively (Fig. 3A). Most SSRs were distributed in LSC (Fig. 3B).

In dispersed repeat detection, 818 forward, 820 palindromic, 11 reverse, and 12 complementary repeats were found, totaling 1661. Among these, eight displayed reverse repeats, and eight displayed complementary repeats. *L. chrysocomum* had the most reverse repeats, *L. tenellum* had the most complementary repeats, and *L. tetragonum* lacked both repeats. (Fig. 3C). The vast majority of these repeats were between 30 and 39 bp, constituting 56.03% of the total repeats. Less than 15.04% of the repeats were 50 bp or longer, with *L. tenellum* having the highest number (38) of such long repeats. In addition, 571 tandem repeat sequences were identified in 12 *Limonium* species. The average number of tandem repeat patterns for each chloroplast genome was 47.58 ± 1.37, ranging from 39 to 54 (Fig. 3C).

Analysis of relative synonymous codon usage (RSCU) across the 12 *Limonium* species showed consistency among the different species, except for a lack of glycine (Gly) in *L. bicolor* (Fig. 3D). Leucine (Leu) is the most frequently used amino acid (27,350), followed by cysteine (Cys) (2,721). The chloroplast genomes of *Limonium* species contained 64 codons for 20 amino acids. Leucine had the highest number of codons and the lowest number of codons. A total of 30 codons displayed an RSCU value >1, with 29 ending in A/U, and 32 codons had an RSCU value ≥ 1, with 29 ending in G/C. This indicated a preference for A/U-ending codons over G/C-ending codons. The

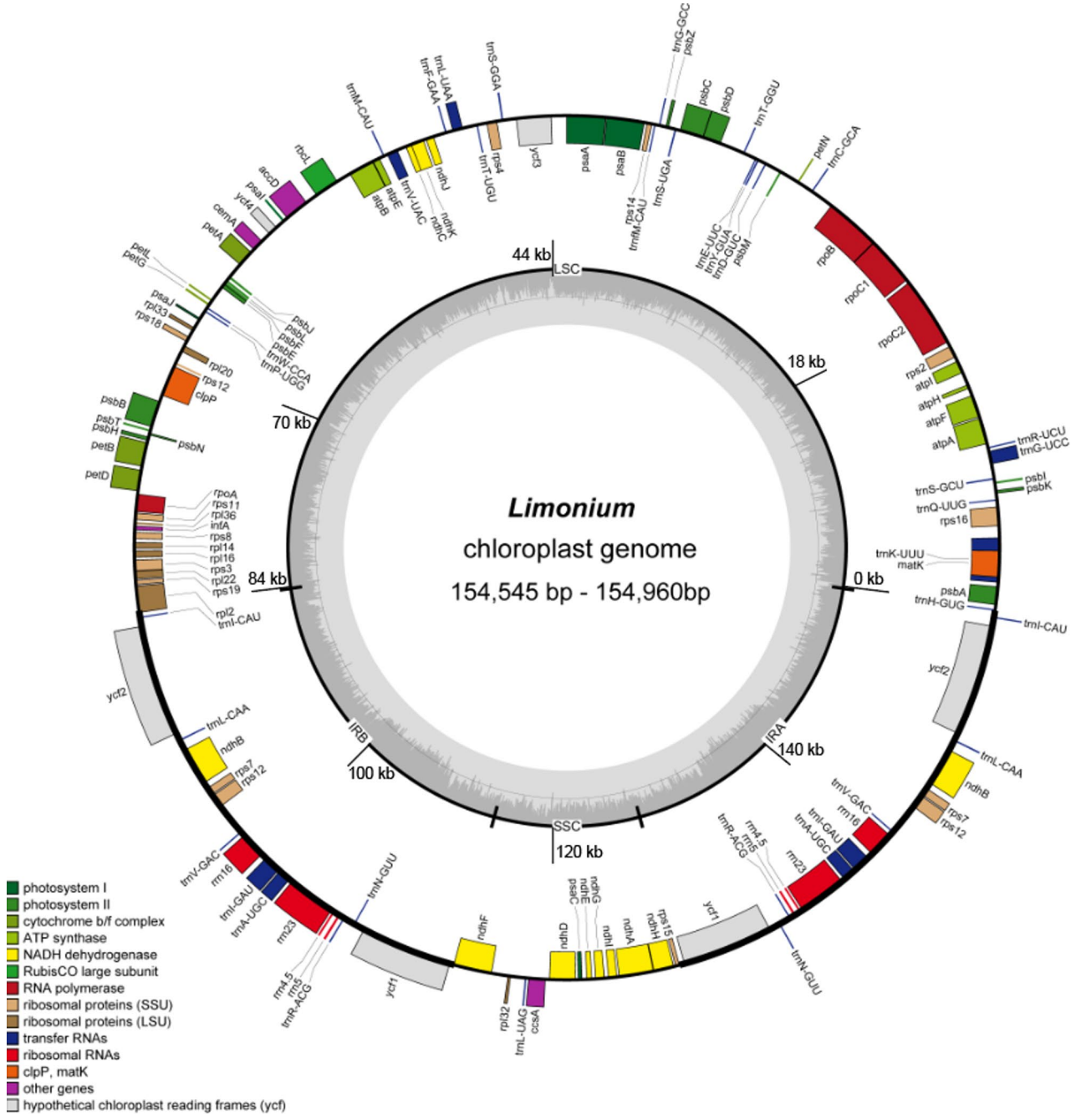

Fig. 1 Chloroplast genomic map of *Limonium* species. Genes inside the circle are transcribed clockwise, whereas those outside are transcribed counterclockwise. The dark and light gray area in the inner circle corresponds to the GC content and AT content, respectively

highest and lowest RSCU values were observed for UUA and CUC, respectively, both of which encode leucine. The effective number of codons (ENC) generally varied between 20 and 61, with lower ENC values signifying a more pronounced preference for codon usage as opposed to a random selection [36, 37]. Our results revealed an ENC range of 48.95 (*L. kaschgaricum*, *L. chrysocomum*) to 49.12 (*L. otolepis*, *L. myrianthum*) across the 12 *Limonium* chloroplast genomes, with an average of 49.01.

Phylogenetic analysis

Based on the chloroplast genomic data of the 12 species, we generated two phylogenetic trees using Bayesian inference (BI) and maximum likelihood (ML) approaches. The topologies of the two trees displayed similarities (Fig. 4; Fig S1). The two resulting phylogenetic trees showed that *L. suffruticosum* and *L. gmelinii* were separate from the other ten *Limonium* species. Almost all the nodes received powerful support (PP/BS = 1/100).

Table 1 Summary of chloroplast genomes of *Limonium* species

Species	Genome size (bp)	LSC (bp)	IR (bp)	SSC (bp)	Number of genes	Protein coding gene	tRNA	rRNA	GC content (%)	LSC (%)	IR (%)	SSC (%)
Limonium aureum	154,545	84,427	28,569	12,980	128	83(5)	37(7)	8(4)	37.1	35.4	40.8	31.5
Limonium bicolor	154,617	84,542	28,555	12,965	128	83(5)	37(7)	8(4)	37	35.3	40.9	31.5
Limonium chrysocomum	154,836	84,568	28,584	13,100	128	83(5)	37(7)	8(4)	37	35.3	40.8	31.4
Limonium franchetii	154,673	84,578	28,559	12,977	128	83(5)	37(7)	8(4)	37	35.3	40.9	31.5
Limonium gmelinii	154,711	84,488	28,535	13,153	128	83(5)	37(7)	8(4)	36.9	35.2	40.8	31.2
Limonium kaschgaricum	154,960	84,616	28,592	13,160	128	83(5)	37(7)	8(4)	37	35.3	40.9	31.3
Limonium leptolobum	154,879	84,620	28,574	13,111	128	83(5)	37(7)	8(4)	37	35.3	40.8	31.3
Limonium myrianthum	154,890	84,651	28,623	12,993	128	83(5)	37(7)	8(4)	37	35.3	40.8	31.5
Limonium otolepis	154,916	84,638	28,623	13,032	128	83(5)	37(7)	8(4)	37	35.3	40.8	31.4
Limonium suffruticosum	154,649	84,439	28,545	13,120	128	83(5)	37(7)	8(4)	37	35.3	40.8	31.3
Limonium tenellum	150,515	84,634	21,064	23,753	128	83(5)	37(7)	8(4)	37.1	35.2	42.8	33.6
Limonium tetragonum	154,691	84,568	28,563	12,997	128	83(5)	37(7)	8(4)	37	35.3	40.9	31.5

The 12 species of *Limonium* were classified into three sections: Sect. *Sarcophyllum*, Sect. *Limonium* and Sect. *Platyhymenium*. Within this phylogenomic framework, representatives of Sect. *Platyhymenium* formed a well-supported monophyletic cluster (BS = 100), with *L. chrysocomum* being a sister to *L. leptolobum* (BS = 100), *L. tetragonum* to *L. franchetii*, and *L. kaschgaricum* and *L. tenellum* identified as sequential sister taxa (BS = 100) to the clade comprising *L. chrysocomum* and *L. leptolobum*. In the Sect. *Limonium*, *L. myrianthum* and *L. otolepis* are sister branches, whereas *L. gmelinii* and *L. suffruticosum* belong to Sect. *Sarcophyllum* are sister branches that formed a monophyletic group (BS = 100).

Discussion

We sequenced eight *Limonium* species in the present study. These chloroplast genomes maintained a consistent size and structure, measuring between 154,545 and 154,960 bp. Little variation has been observed in the IR and SSC regions, yet these evolutionary changes can influence chloroplast genomes sizes [38–40]. Previous studies have shown that IR expansion in Plumbaginaceae, which typically accommodates the *ycf1* gene, is absent in *Limonium* [31]. However, our results indicated that the IR has been expanded to include *ycf1* in all *Limonium* chloroplast genomes, with the exception of *L. tenellum*, which is consistent with the conclusions of Darshetkar et al. [35]. The annotation provided for *L. tenellum* (MK397871) shows a single copy of *ycf1* gene. Our re-annotation yielded resulted consistent with other *Limonium*, revealing that *ycf1* gene had two copies in *L. tenellum*. In terms of genome size, *L. tenellum* is only about 150 kb, about 4 kb less than most *Limonium* species. What's more, *L. tenellum* underwent massive variation from 104 kb to 116 kb. Therefore, *L. tenellum* could be attributed to assembly and annotation errors [35].

The loss of introns is often associated with genome simplification and adaptive evolution and is generally thought to involve a reverse-transcriptase-mediated mechanism [41]. Our study identified the loss of the *rpl16* intron in all eight sequenced *Limonium* species. While this finding is consistent with previous reports in three genera of Plumbaginaceae, including *Limonium* [35, 42–44], merely confirming its occurrence. The ubiquitous absence of this intron across the sampled species strongly suggests that the loss event was not random but occurred once in a common ancestor of the studied *Limonium* lineage. This makes the *rpl16* intron loss a potential shared derived trait for this group, providing a valuable molecular character for defining this clade within the genus [42]. The loss of introns in the *rpl16* gene may represent a neutral evolutionary event achieved through reverse transcription mechanisms, possessing phylogenetic marker

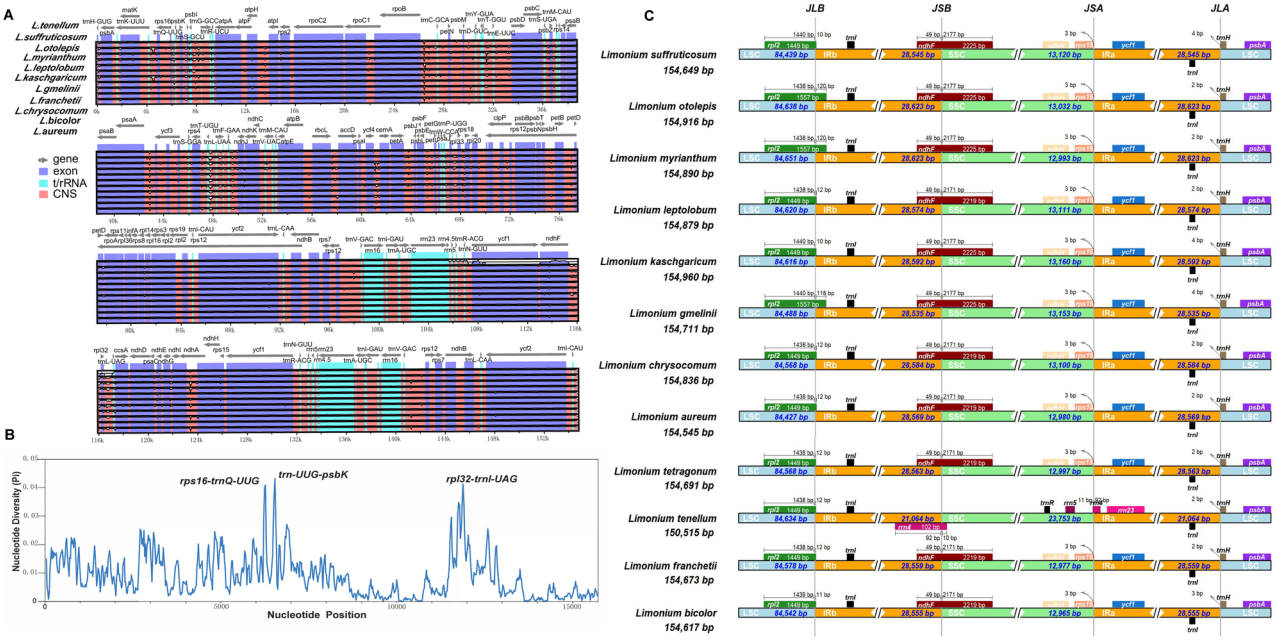

Fig. 2 Comparative genomic analysis in chloroplast genomes of *Limonium* species. **A** Sequence alignment of the chloroplast genomes in 12 species. *L. tetragonum* (MW085088) was used as a reference. The y-axis indicates the percent identity between 50% and 100%. The white peaks indicate the regions with sequence variation among *Limonium*. The purple, blue, pink and gray bars correspond to exons, untranslated regions, noncoding sequences and mRNAs, respectively. **B** Comparison of nucleotide variability of common genes in 11 species. Window length: 600 bp; step size: 200 bp. x-axis: Range of sequence length; y-axis: Pi value. **C** LSC, SSC and IR boundaries of the chloroplast genomes in 12 species. JLB: junction of LSC and IRb; JSB: junction of SSC and IRb; JSA: junction of SSC and IRa; JLA: junction of LSC and IRa. Boxes above and below the mainline indicate the adjacent border genes. The gaps between the genes and the boundaries are indicated by the base lengths

value. It may also confer a slight selective advantage in specific contexts by reducing splicing burden [45].

The *rpl23* gene encodes a ribosomal protein, and its functional transfer to the nucleus followed by inactivation of the chloroplast copy is a well-documented phenomenon in plant evolution [26, 32, 46–49]. In this study, the *rpl23* gene was interrupted by encountering a stop codon in advance, which might have undergone pseudogenization. In Caryophyllales, pseudogenization of the *rpl23* gene was widespread with at least 11 independent losses [32]. Additionally, *rpl23* is interrupted by internal stop codons in the family Polygonaceae, a sister family of Plumbaginaceae, indicating its pseudogene status [46]. Our results further validated these previous findings. We proposed that the functional *rpl23* gene might have been relocated to the nucleus because this phenomenon is not uncommon, as numerous plastid genes in flowering plants are known to have experienced such a transfer [26, 47, 48], and the same results were predicted in the members of Plumbaginaceae [32]. The pseudogenization of *rpl23* likely represents a case of functional redundancy elimination and genomic optimization [49]. Once a functional, nuclear-encoded copy of *rpl23* was established and the protein product could be successfully imported back into the chloroplast, the maintenance of the original chloroplast copy became superfluous. Mutations accumulating in the chloroplast *rpl23* would then be neutral in effect, leading to its pseudogenization [50, 51]. This process relaxes the evolutionary constraints on the chloroplast copy, allowing it to evolve at a faster rate. Therefore, the pseudogenized *rpl23* in *Limonium* is not merely a broken gene but is a marker of a completed evolutionary transition. This characteristic could be exploited in future phylogenetic studies, as its presence or absence and its specific mutation patterns might provide valuable synapomorphic signals for resolving relationships at specific taxonomic levels within Plumbaginaceae [50, 51].

Repetitive sequences within chloroplast genomes play a vital role in genome rearrangement and genetic diversity [52]. In the present study, we identified 1,661 repeat sequences, with regions containing these repeats serving as significant hotspots for genomic reconfiguration. These sequences offer valuable insights into the evolutionary history and genetic divergence of plant species [53–55]. Additionally, they might yield a wealth of data for population genetics and phylogenetic analyses [53, 56, 57]. Of the SSRs, A/T bp represented the highest proportion, whereas C and G nucleotides had the lowest. This pattern also is found in other medicinal plants such as *Alpinia* [58], *Aconitum* [59], and *Scrophularia* [60]. SSRs within the chloroplast genome are commonly employed as genetic markers in population and evolutionary studies [61–63].

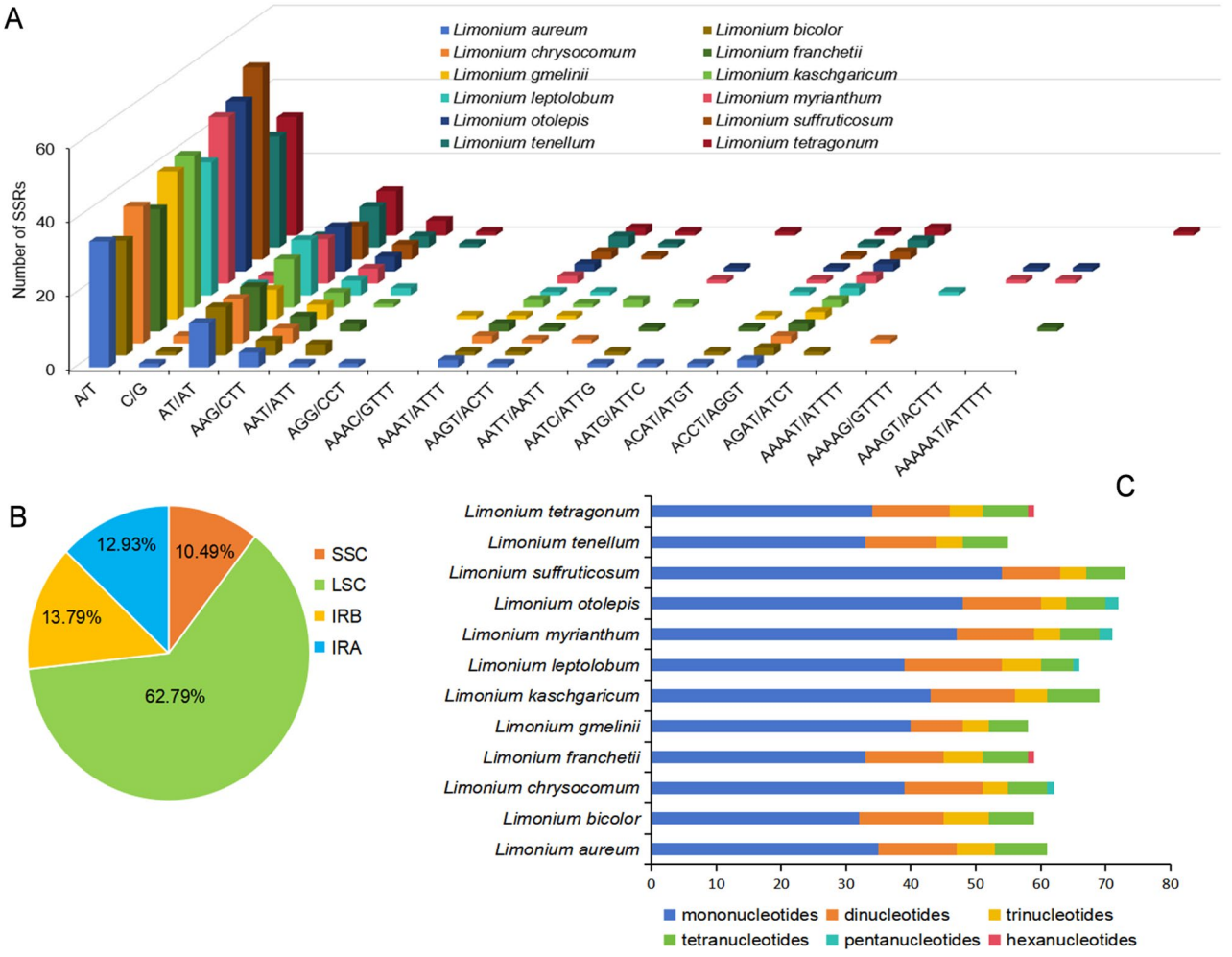

Fig. 3 Repeat sequences and codon usage bias analysis among 12 chloroplast genomes. **A** Numbers of SSRs detected. **B** SSRs locus distribution (including LSC, SSC, IRa and IRb).**C** Frequencies of all repeat sequence. **D** Heatmap of relative synonymous codon usage (RSCU) values. The four gray squares represent the lack of corresponding codons (Glycine) in *L. bicolor*

DNA sequences from various organisms clearly demonstrate that synonymous codons for a particular amino acid are not uniformly distributed, despite the fact that these codons should theoretically be used equally in terms of their impact on protein structure [64]. Overall, genomic GC content has been identified as the most influential factor affecting variation in codon usage among different species [65]. In the present study, the mean GC content of codons in the chloroplast genomes was 37.6%. The GC3 content ranged from 25.5% (*L. gmelinii*) to 25.7% (*L. aureum*, *L. franchetii*, *L. myrianthum*, *L. otolepis*, *L. suffruticosum*, *L. tenellum*, and *L. tetragonum*), all below 50%. Of the total codons, 30 had a RSCU value exceeding 1, with the exception of one, all ending in A/U, suggesting a bias towards A/U-ending codons in chloroplast genomes, which aligns with most studies on chloroplast genomes [60, 66–69].

The phylogenetic position of *Limonium* species studied using chloroplast genomes was consistent with previous studies based on chloroplast genomic fragments [1, 23]. Phylogenetic analysis showed that all *Limonium* species formed a large branch with strong support. *L. suffruticosum* and *L. gmelinii* were separated from the other ten *Limonium* species. Eight *Limonium* spp. from Sect. *Platyhymenium* has a relatively larger calyx and longer tubular corolla; *L. tenellum*, *L. franchetii*, and *L. kaschgaricum* possess light purple-red corolla and the other five species have orange-yellow or yellow corolla. The species of Sect. *Sarcophyllum* and Sect. *Limonium* has a relatively small calyx and a short tubular corolla. Both possess corollas ranging from light purple to blue-purple.

The phylogenetic tree reveals that the different sections of *Limonium* examined in this study (Sect. *Platythymenium*, Sect. *Limonium*, and Sect. *Sarcophyllum*) form well-resolved clades, with most nodes showing strong support (BS = 100). Regarding the phylogenetic placement of these taxa, multiple species of Sect. *Platythymenium* constitute a highly supported monophyletic

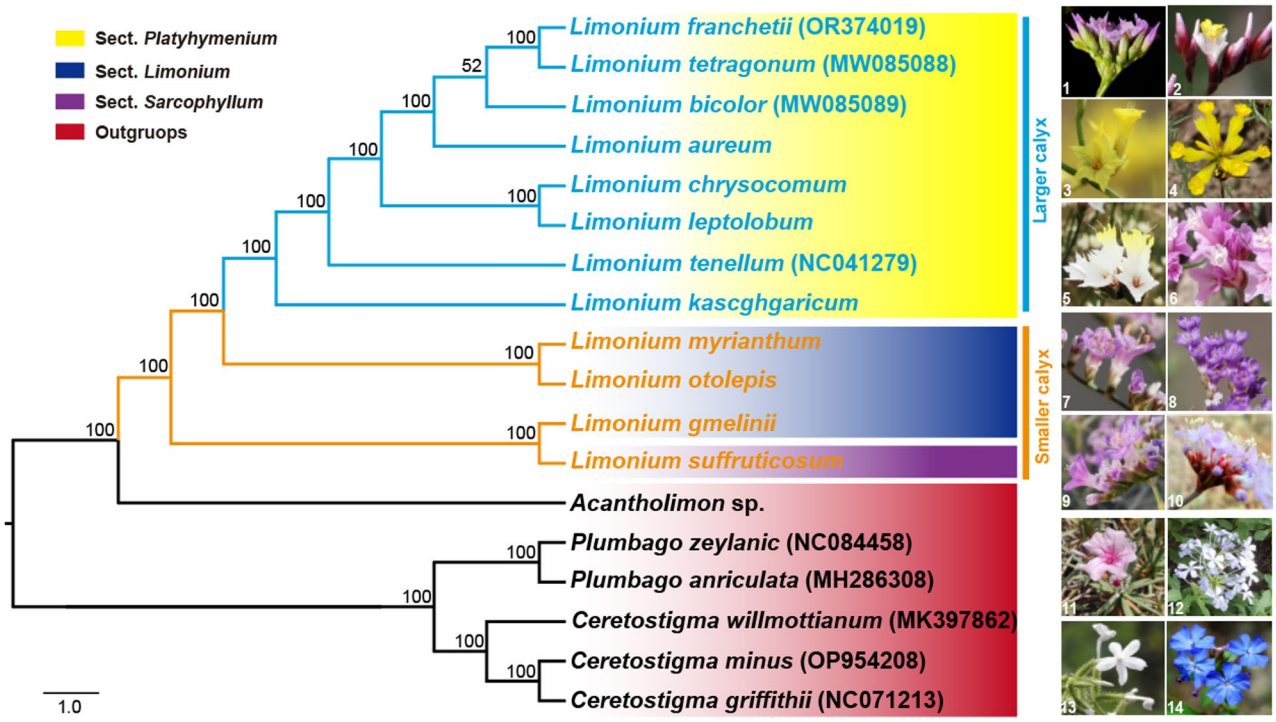

Fig. 4 Maximum Likelihood tree based on *Limonium* complete chloroplast genomic sequences. Six species were selected as outgroups from other genera of the Plumbaginaceae. The yellow section represents Sect. *Platyhymenium*, the blue section represents Sect. *Limonium*, the purple section represents Sect. *Sarcophyllum*, and the red section represents the outgroups. The node numbers indicate maximum likelihood bootstrap support values. Representative plants of *Limonium* and Plumbaginaceae are shown on the right: 1 *L. franchetii* (image courtesy of Hong Zhao), 2 *L. bicolor*, 3 *L. aureum*, 4 *L. chrysocomum*, 5 *L. leptolobum*, 6 *L. kascghgaricum*, 7 *L. myrianthum*, 8 *L. otolepis*, 9 *L. gmelinii*, 10 *L. suffruticosum*, 11 *Acantholimon* sp. 12 *P. anriculata* 13 *P. zeylanic* 14 *C. willmottianum*; 3 and 6-8 photos by Aiqin Zhang; 4, 5, 9 and 10 photos by Fangfang Jiao; the photo of other plants is from https://powo.science.kew.org/

clade, consistent with previous findings that the genus can be divided into distinct subgenera and lineages [1]. This indicates that chloroplast genome data are effective in resolving phylogenetic relationships among species within this section. Furthermore, Sect. *Limonium* and Sect. *Sarcophyllum* each form relatively independent clades, supporting the view that different sections of *Limonium* have undergone distinct evolutionary trajectories [20]. This suggests that species belonging to these sections in Xinjiang, China, have accumulated substantial genetic divergence over long-term evolution, resulting in stable phylogenetic relationships. In addition, outgroup taxa, including *Acantholimon* sp., *Plumbago*, and *Ceratostigma*, are clearly separated from the *Limonium* species. This observation validates previous conclusions regarding the delineation between *Limonium* and its related genera [18], confirming the reliability of the present phylogenetic analysis at the generic level. From a biogeographic perspective, Xinjiang (Northwest of China) is located at the eastern margin of the Central Asian–Irano–Turanian region. The branching patterns of the different *Limonium* sections in the phylogeny suggest the presence of multiple evolutionary lineages in this region with historical connections to adjacent areas. It is plausible that the *Limonium* species in Xinjiang have a complex origin and dispersal history, with some lineages potentially sharing common ancestors with those from the Irano–Turanian or Mediterranean regions [1, 18, 20]. These lineages may have gradually diverged due to long-term geographic isolation and environmental adaptation. Moreover, certain species in Sect. *Platythymenium* may have accumulated adaptive mutations in their chloroplast genomes in response to the arid and saline habitats specific to Xinjiang, China. This aligns with the proposal by Darshetkar et al. [35] that positively selected genes in the chloroplast genome may be associated with stress adaptation, implying that these species have undergone adaptive changes in chloroplast genes to enhance their survival in extreme environments.

Previous studies have reported chloroplast genomic data focused on a single (or four) *Limonium* species and positioned these species in close relation to those of the *Plumbago* genus [32, 34, 35, 58]. In this phylogenetic study, we included 12 species for the first time and conducted a comparative analysis of their chloroplast genomes. However, given the substantial diversity within the *Limonium* genus, which comprises approximately 600 species [1], the analysis of additional *Limonium* chloroplast genomes could provide greater insight into intrageneric relationships and trait evolution. It has

been reported that nuclear-chloroplast conflicts exist in Mediterranean lineages [1], but broader sampling is still required to determine evolutionary patterns between nuclei and chloroplasts in the Asian-Iranian-Turanian region.

Conclusions

The newly assembled complete chloroplast genomes of the eight *Limonium* species exhibited a relatively conserved chloroplast genomic structure consisting of 78 protein-coding genes, 30 tRNA genes, and four rRNA genes. The *rpl16* intron was lost and the *rpl23* gene underwent pseudogenization. A comparative analysis revealed a minor expansion in the IR regions of the complete chloroplast genomes, which may have implications for genome stability and evolution in *Limonium* species. We identified 55–73 SSRs that provide valuable DNA markers for population genetics and molecular phylogenetic studies. High nucleotide diversity was detected in three regions: *rps16-trnQ-UUG*, *rpl32-trnL-UAG*, and *rps16-trnQ-UUG*, suggesting their potential utility as DNA barcodes for identifying species within the genus. Comparative genomic analysis revealed a preference for A/U-ending codons in the examined species. The results of the phylogenetic tree support that species belonging to these sections in Xinjiang, China, have accumulated substantial genetic divergence over long-term evolution, resulting in stable phylogenetic relationships. These findings will facilitate further investigation of the taxonomy, phylogenetic evolution, and exploration and utilization of *Limonium* genus.

Methods

Taxon sampling, DNA extraction, and sequencing

We collected samples from eight *Limonium* species. All identifications of the specimens were verified by the authors. The collected materials were dried using silica gel, and the voucher specimens have been deposited in Xinjiang University Herbarium; detailed sample data and distribution are provided (Fig S2; Table S2).

Genomic DNA was extracted from moderately dried tissues using a modified Cetyltrimethylammonium Bromide (CTAB) method [70]. Genomic DNA was sheared to an average size of 350 base pairs (bp) by Covaris sonication. The resulting fragments underwent end polishing and addition of an A-tail and were equipped with full-length adapters suitable for Illumina sequencing, followed by PCR amplification. The PCR products were purified using the AMPure XP system (Beverly, MA, USA). The quality of the sequencing libraries was evaluated on an Agilent 5400 system (Agilent, USA), and their quantity was determined as 1.5 nM using qPCR. Qualified libraries were consolidated and subjected to sequencing on Illumina platforms employing the PE150 strategy at Novogene Bioinformatics Technology Co., Ltd. (Beijing, China). In addition, we downloaded the complete chloroplast sequences of the four species from the NCBI database for comparative genomic analysis with the following accession numbers: OR374019, MW085088, MW085089, and NC_041279.

Chloroplast genomes assembly and annotation

Chloroplast genomes were assembled in GetOrganelle v1.7.7.0 [71] using default parameters. A comparison of the two assembled sequences were conducted using Gepard v1.40, and the sequence that aligned in the same direction as the SSC region in the reference sequence was selected [72]. Subsequently, the assembled sequences were annotated using CPGview [73] and the Geseq platform [74], using *L. tetragonum* (MW085088), as references. The annotation results were checked and manually adjusted using Geneious Prime 2024.05 [75]. The chloroplast genome map was visualized using the Geseq platform [74]. The complete chloroplast genomes of the eight *Limonium* species were submitted to GenBank (Table S2).

Comparative Chloroplast genomic analysis

We compared the chloroplast genomes across all published *Limonium* species using the online genome comparison tool mVISTA [76], except for *L. sinense* owing to ambiguities observed in the assembly. Subsequently, protein-coding genes (PCGs) from the chloroplast genomes were extracted using Geneious Prime and aligned using MAFFT v7.505 [77, 78]. Nucleotide diversity analysis was performed using DnaSP v6.0 [79]. Finally, we used IRscope to detect expansions or contractions in the IR regions.

Repeat sequences and codon usage bias analysis

We identified dispersed, simple, and tandem repeat sequences within the chloroplast genomes of 12 *Limonium* species. Dispersed repeats were identified using the REPuter online tool [80], which included forward, reverse, palindromic, and complementary repeats, were identified using the REPuter online tool. The parameters used were a Hamming distance of 3, a minimum repeat size of 30 bp, and a maximum repeat count of 5,000. SSRs were predicted using the online tool MISA [81]. Tandem repeats were identified using Tandem Repeats Finder with the default settings.

RSCU and GC content at the third position (GC3) were calculated using CodonW v1.4.2 and the results were visualized in RStudio 2023.12.0 + 369 using the *pheatmap* package [82].

Phylogenetic analysis

We used chloroplast genomes to construct ML and BI trees to determine the phylogenetic relationships within the 12 species with *Ceratostigma willmottianum*, *C. griffithii*, *C. minus*, *Plumbago auriculate*, *P. zeylanica*, and *Acantholiton* sp. were selected as outgroups. All chloroplast genomic sequences were aligned using MAFFT v7.490 default parameter. Subsequently, Trimal v1.2rev57 (-gt 0.2 -st 0.001 -cons 85) software was employed to trim these aligned sequences to obtain a supermatrix [83]. The ML tree was determined using IQ-tree v2.2.0 (-m MFP) based on the best fit model GTR + F + R5 [84], with the bootstrap (BS) parameter set to 1000. BI tree was constructed using BI in MrBayes v3.2.7a based on the best fit model GTR + F + I + G4 [85], implementing a partition model that involved two simultaneous runs and a total of 2,000,000 generations, with the first 25% of the samples set aside as burn-in to account for initial instability in Markov chain Monte Carlo (MCMC) simulations.

Abbreviations

ITS	Inter transcribed spacer
kb	kilobase pairs
LSC	Large single-copy region
SSC	Small single-copy
IR	Inverted repeat
rRNA	Ribosomal RNA
tRNA	Transfer RNA
AT	Adenine–thymine
GC	Guanine-cytosine
Pi	Nucleotide diversity
SC	Single-copy regions
SSRs	Simple sequence repeats
RSCU	Relative synonymous codon usage
Gly	Glycine
Leu	Leucine
Cys	Cysteine
ENC	Effective number of codons
BI	Bayesian inference
ML	Maximum likelihood
CTAB	Cetyltrimethylammonium bromide
PCGs	Protein-coding genes
MCMC	Markov chain Monte Carlo

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-07913-9.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Acknowledgements

This work was supported by National Natural Science Foundation of China [32360308, 32360058] and the Government of Xinjiang Uygur Autonomous Region of China [XJ2023G029, XJEDU2021I006]. We would like to thank Professor Hong Zhao from Shandong University for providing image.

Authors' contributions

All authors contributed to the study conception and design. Conceptualization, data curation, formal analysis, visualization, investigation, methodology, software and writing — original draft preparation, FFJ; investigation, validation and formal analysis, XHY and JZ; writing—review and editing, GY; resources, supervision and project administration, AQZ; writing—review and editing, formal analysis, visualization, funding acquisition, XMT. All authors read and approved the final manuscript.

Funding

This work was supported by National Natural Science Foundation of China [32360308, 32360058]; Government of Xinjiang Uygur Autonomous Region of China [XJ2023G029, XJEDU2021I006].

Data availability

Sequence data that support the findings of this study have been deposited in the National Center for Biotechnology Information (NCBI) with accession number PQ740888-PQ740895. The voucher number and collector of the species are as follow, ***Limonium aureum*** (ZAQ20180515, Aiqin Zhang), ***L. chrysocomum*** (JFF20220530, Fangfang Jiao), ***L. gmelinii*** (JFF20220817, Fangfang Jiao), ***L. kaschgaricum*** (ZAQ20180517, Aiqin Zhang), ***L. leptolobum*** (JFF20220523, Fangfang Jiao), ***L. myrianthum*** (JFF20210615, Fangfang Jiao), ***L. otolepis*** (JFF20210713, Fangfang Jiao), ***L. suffruticosum*** (JFF20220823, Fangfang Jiao).

Declarations

Ethics approval and consent to participate

This research did not involve approval from an ethics review board as it did not entail the use of human subjects, animal subjects, or any materials and methods requiring ethical approval. Therefore, an ethics statement is not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 5 December 2024 / Accepted: 4 December 2025
Published online: 18 December 2025

References

1. Koutroumpa K, Theodoridis S, Warren BH, Jiménez A, Celep F, Doğan M, et al. An expanded molecular phylogeny of Plumbaginaceae, with emphasis on *Limonium* (sea lavenders): taxonomic implications and biogeographic considerations. Ecol Evol. 2018;8:12397–424.
2. Doğan M, Akaydı G, Erdal J. Phytogeography, ecology and conservation of the genus *Limonium* (Plumbaginaceae) in Turkey and a taxonomic revision. Plant Syst Evol. 2020;306:89.
3. Peng ZX, Rudolf VK. Plumbaginaceae. In: Wu ZY, Peter HR, editors. Flora of China. Volume 15. Beijing, China: Science; 1996.
4. Kubitzki K. Plumbaginaceae. In: Kubitzki K, Rohwer JG, Bittrich V, editors. The families and genera of vascular plants. Volume 2. Berlin, Germany: Springer; 1993.
5. Caperta AD, Róis AS, Teixeira G, Garcia-Caparros P, Flowers TJ. Secretory structures in plants: lessons from the Plumbaginaceae on their origin, evolution and roles in stress tolerance. Plant Cell Environ. 2020;43:2912–31.
6. Yuan F, Wang X, Zhao B, Xu X, Shi M, Leng B, et al. The genome of the recretohalophyte *Limonium bicolor* provides insights into salt gland development and salinity adaptation during terrestrial evolution. Mol Plant. 2022;15:1024–44.
7. González-Oreng S, Grigore MN, Boscaiu M, Vicente O. Constitutive and induced salt tolerance mechanisms and potential uses of *Limonium* Mill. species. Agronomy. 2021;11:413.
8. Eren Y, Özata A. Determination of mutagenic and cytotoxic effects of *Limonium globuliferum* aqueous extracts by *Allium*, Ames, and MTT tests. Rev Bras Farmacogn. 2014;24(1):51–9.
9. Medini F, Ksouri R, Falleh H, Megdiche W, Trabelsi N, Abdelly C. Effects of physiological stage and solvent on polyphenol composition, antioxidant and antimicrobial activities of *Limonium densiflorum*. J Med Plant Res. 2011;5(31):6719–30.

10. Hamadou MH, Kerkatou M, Gatto P, Pancher M, Bisio A, Inga A, et al. Apigenin rich-*Limonium duriusculum* (de Girard) Kuntze promotes apoptosis in HCT116 cancer cells. Nat Prod Res. 2021;35(17):2910–4.
11. Lee YG, Song MY, Cho H, Jin JS, Park BH, Bae EJ. *Limonium tetragonum* promotes running endurance in mice through mitochondrial biogenesis and oxidative fiber formation. Nutrients. 2022;14:3904.
12. Gancedo NC, Isolani R, de Oliveiran NC, Nakamura CV, de Medeiros Araújo DC, Sanches ACC, et al. Chemical constituents, anticancer and anti-proliferative potential of *Limonium* species: a systematic review. Pharmaceuticals. 2023;16:293.
13. Sanna C, Rigano D, Corona A, Piano D, Formisano C, Farci D, Franzini G, Ballero M, Hianese G, Tramontano E, et al. Dual HIV-1 reverse transcriptase and integrase inhibitors from *Limonium morisianum* Arrigoni, an endemic species of Sardinia (Italy). Nat Prod Res. 2018;54:1–6.
14. Mori S, Yahata M, Kuwahara A, Shirono Y, Ueno Y, Hatanaka M, et al. Morphological characterization of tetraploids of *Limonium sinuatum* (L.) Mill. produced by *Oryzalin* treatment of seeds. Horticulturae. 2021;7:248.
15. Xu XJ, Zhou YL, Mi P, Wang BS, Yuan F. Salt–tolerance screening in *Limonium sinuatum* varieties with different flower colors. Sci Rep. 2021;11:14562.
16. Cowan RMJ, Ingrouille M, Lledó MD. The taxonomic treatment of agamosperms in the genus *Limonium* Mill. (Plumbaginaceae). Folia Geobot. 1998;33(3):353–66.
17. Palacios C, Rosselló JA, González-Candelas F. Study of the evolutionary relationships among *Limonium* species (Plumbaginaceae) using nuclear and cytoplasmic molecular markers. Mol Phylogenet Evol. 2000;14(2):232–49.
18. Akhani H, Malekmohammadi M, Mahdavi P, Gharibiyan A, Chase MW. Phylogenetics of the *Irano-Turanian* taxa of *Limonium* (Plumbaginaceae) based on ITS nrDNA sequences and leaf anatomy provides evidence for species delimitation and relationships of lineages. Bot J Linn Soc. 2013;171(3):519–50.
19. Róis AS, Sádio F, Paulo OS, Teixeira G, Paes AP, Espírito-Santo D, et al. Phylogeography and modes of reproduction in diploid and tetraploid halophytes of *Limonium* species (Plumbaginaceae): evidence for a pattern of geographical parthenogenesis. Ann Bot. 2016;117(1):37–50.
20. Malekmohammadi M, Akhani H, Borsch T. Phylogenetic relationships of *Limonium* (Plumbaginaceae) inferred from multiple chloroplast and nuclear loci. Taxon. 2017;66(5):1128–46.
21. Lledó MD, Crespo MB, Fay MF, Chase MW. Molecular phylogenetics of *Limonium* and related genera (Plumbaginaceae): biogeographical and systematic implications. Am J Bot. 2005;92(7):1189–98.
22. Lledó MD, Karis PO, Crespo MB, Fay MF, Chase MW. Endemism and evolution in Macronesian and mediterranean *Limonium* taxa. In: Bramwell D, Caujapé-Castells J, editors. The biology of Island floras. Cambridge, USA: Cambridge University Press; 2011.
23. Koutroumpa K, Warren BH, Theodoridis S, Coiro M, Romeiras MM, Jiménez A, et al. Geo-climatic changes and apomixis as major drivers of diversification in the mediterranean sea lavenders (*Limonium* Mill). Front Plant Sci. 2021;11:612258.
24. Jansen RK, Ruhlman TA. Plastid genomes of seed plants. In: Bock R, Knoop V, editors. Genomics of chloroplasts and mitochondria, advances in photosynthesis and respiration. Dordrecht, Switzerland: Springer; 2012. pp. 103–26.
25. Wicke S, Schneeweiss GM, Müller KF, Quandt D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant Mol Biol. 2011;76:273–97.
26. Daniell H, Lin CS, Yu M, Chang WJ. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biol. 2016;17:134.
27. Wang J, Kan SL, Liao XZ, Zhou JW, Tembrock LR, Daniell H, et al. Plant organellar genomes: much done, much more to do. Trends Plant Sci. 2024;29(7):754–69.
28. Gompert Z, Mock KE. Detection of individual ploidy levels with genotyping-by-sequencing (GBS) analysis. Mol Ecol Resour. 2017;17(6):1156–67.
29. Thomas GWC, Ather SH, Hahn MW. Gene-tree reconciliation with MUL-Trees to resolve polyploidy events. Syst Biol. 2017;66(6):1007–18.
30. McKain MR, Johnson MG, Uribe-Convers S, Eaton D, Yang Y. Practical considerations for plant phylogenomics. Appl Plant Sci. 2018;6(3):e1038.
31. Pina–Martins F, Caperta AD, Conceição SIR, Nunes VL, MarquesI, Paulo OS. A first look at sea–lavenders genomics – can genome wide SNP information tip the scales of controversy in the *Limonium vulgare* species complex? BMC Plant Biol. 2023;23:34.
32. Yao G, Jin JJ, Li HT, Yang JB, Mandala VS, Croley M, et al. Plastid phylogenomic insights into the evolution of Caryophyllales. Mol Phylogenet Evol. 2019;134:74–86.
33. Li J, Xu B, Yang Q, Wang T, Zhu Q, Lin Y, et al. The complete chloroplast genome sequence of *Limonium sinense* (Plumbaginaceae). Mitochondr DNA B Resour. 2020;5(1):556–7.
34. Zhang X, Xu Y, Liu X. Complete plastome sequence of *Limonium aureum*, a medicinal and ornamental species in China. Mitochondr DNA B Resour. 2020;5(1):333–4.
35. Darshetkar AM, Maurya S, Lee C, Bazarragchaa B, Batdelger G, Janchiv A, et al. Plastome analysis unveils inverted repeat (IR) expansion and positive selection in sea lavenders (*Limonium*, Plumbaginaceae, Limonioideae, Limonieae). PhytoKeys. 2021;175:89–107.
36. Wu X, Wu S, Ren D, Zhu Y, He F. The analysis method and progress in the study of codon bias. Hereditas. 2007;29:420–6.
37. Li Q, Luo Y, Sha A, Xiao W, Xiong Z, Chen X, et al. Analysis of synonymous codon usage patterns in mitochondrial genomes of nine *Amanit*a species. Front Microbiol. 2023;14:1134228.
38. Zhu A, Guo W, Gupta S, Fan W, Mower JP. Evolutionary dynamics of the plastid inverted repeat: the effects of expansion, contraction, and loss on substitution rates. New Phytol. 2016;209(4):1747–56.
39. Darshetkar AM, Datar MN, Tamhankar S, Li P, Choudhary RK. Understanding evolution in poales: insights from eriocaulaceae plastome. PLoS One. 2019;14(8):e0221423.
40. Maurya S, Darshetkar AM, Datar MN, Tamhankar S, Li P, Choudhary RK. Plastome data provide insights into intra and interspecific diversity and *Ndh* gene loss in capparis (Capparaceae). Phytotaxa. 2020;432(2):206–20.
41. Downie SR, Olmstead RG, Zurawski G, Soltis DE, Soltis PS, Watson JC, et al. Six independent losses of the chloroplast DNA *rpl2* intron in dicotyledons: molecular and phylogenetic implications. Evolution. 1991;45(5):1245–59.
42. Campagna ML, Downie SR. The intron in chloroplast gene *rpl16* is missing from the flowering plant families Geraniaceae, Goodeniaceae, and Plumbaginaceae. Trans Ill State Acad Sci. 1998;91:1–11.
43. Downie SR, Palmer JD. Use of Chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Soltis PS, Soltis DE, Doyle JJ, editors. Molecular systematics of plants. New York: Chapman and Hall; 1992. pp. 14–35.
44. Downie SR, Palmer JD. A chloroplast DNA phylogeny of the caryophyllales based on structural and inverted repeat restriction site variation. Syst Bot. 1994;19:236–52.
45. Belshaw R, Bensasson D. The rise and falls of introns. Heredity. 2006;96:208–13.
46. Logacheva MD, Samigullin TH, Dhingra A, Penin AA. Comparative chloroplast genomics and phylogenetics of *Fagopyrum esculentum* ssp. ancestrale-a wild ancestor of cultivated buckwheat. BMC Plant Biol. 2008;8(1):e59.
47. Millen RS, Olmstead RG, Adams KL, Palmer JD, Lao NT, Heggie L, et al. Many parallel losses of *infA* from chloroplast DNA during angiosperm evolution with multiple independent transfers to the nucleus. Plant Cell. 2001;13(3):645–58.
48. Jansen RK, Saski C, Lee S-B, Hansen AK, Daniell H. Complete plastid genome sequences of three Rosids (*Castanea*, *Prunus*, *Theobroma*): evidence for at least two independent transfers of *rpl22* to the nucleus. Mol Biol Evol. 2011;28(1):835–47.
49. Leister D. Origin, evolution and genetic effects of nuclear insertions of organelle DNA. Trends Genet. 2005;21(12):655–63.
50. Li WH, Gojobori T, Nei M. Pseudogenes as a paradigm of neutral evolution. Nature. 1981;292(5820):237–9.
51. Pink RC, Carter DRF. Pseudogenes as regulators of biological function. Essays Biochem. 2013;54(1):103–12.
52. Yang Q, Jiang Y, Wang Y, Han R, Liang Z, He Q, et al. SSR loci analysis in transcriptome and molecular marker development in *Polygonatum sibiricum*. Biomed Res Int. 2022;2022(1):4237913.
53. Zong D, Gan P, Zhou A, Li J, Xie Z, Duan A, et al. Comparative analysis of the complete chloroplast genomes of seven *Populus* species: insights into alternative female parents of *Populus tomentosa*. PLoS One. 2019;14:e0218455.
54. Sun J, Wang Y, Liu Y, Xu C, Yuan Q, Guo L, et al. Evolutionary and phylogenetic aspects of the chloroplast genome of chaenomeles species. Sci Rep. 2020;10:11466.
55. Chong X, Li Y, Yan M, Wang Y, Li M, Zhou Y, et al. Comparative chloroplast genome analysis of 10 Ilex species and the development of species-specific identification markers. Ind Crops Prod. 2022;187:115408.
56. Gao L, Yi X, Yang Y, Su Y, Wang T. Complete chloroplast genome sequence of a tree fern *Alsophila spinulosa*: insights into evolutionary changes in fern chloroplast genomes. BMC Evol Biol. 2009;9:1–14.

57. Nie X, Lv S, Zhang Y, Du X, Wang L, Biradar SS, et al. Complete chloroplast genome sequence of a major invasive species, Crofton weed (*Ageratina adenophora*). PLoS ONE. 2012;7:e36869.
58. Li D, Zhu G, Xu Y, Ye Y, Liu J. Complete chloroplast genomes of three medicinal Alpinia species: genome organization, comparative analyses and phylogenetic relationships in family Zingiberaceae. Plants. 2020;9:286.
59. Niu Y, Su T, Wu C, Deng J, Yang F. Complete chloroplast genome sequences of the medicinal plant *Aconitum transsectum* (Ranunculaceae): comparative analysis and phylogenetic relationships. BMC Genomics. 2023;24:1–15.
60. Wang X, Guo L, Ding L, Medina L, Wang R, Li P. Comparative plastome analyses and evolutionary relationships of 25 East Asian species within the medicinal plant genus *Scrophularia* (Scrophulariaceae). Front Plant Sci. 2024;15:1439206.
61. Yang Y, Zhou T, Duan D, Yang J, Feng L, Zhao G. Comparative analysis of the complete chloroplast genomes of five *Quercus* species. Front Plant Sci. 2016;7:959.
62. Guo L, Guo S, Xu J, He L, Carlson JE, Hou X. Phylogenetic analysis based on chloroplast genome uncover evolutionary relationship of all the nine species and six cultivars of tree peony. Ind Crops Prod. 2020;53:112567.
63. Chen X, Mao Y, Chai W, Yan K, Liang Z, Xia P. Genome-wide identification and expression analysis of MYB gene family under nitrogen stress in Panax Notoginseng. Protoplasma. 2023;260:189–205.
64. Ikemura T. Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol. 1985;2:13–34.
65. Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias. Nat Rev Genet. 2011;12:32–42.
66. Wu L, Fan P, Cai J, Zang C, Lin Y, Xu Z, et al. Comparative genomics and phylogenomics of the genus *Glycyrrhiza* (Fabaceae) based on chloroplast genomes. Front Pharmacol. 2024;15:1371390.
67. Huang K, Li B, Chen X, Qin C, Zhang X. Comparative and phylogenetic analysis of chloroplast genomes from ten species in *Quercus* section cyclobalanopsis. Front Plant Sci. 2024;15:1430191.
68. Guo L, Wang X, Wang R, Li P. Characterization and comparative analysis of chloroplast genomes of medicinal herb *Scrophularia ningpoensis* and its common adulterants (Scrophulariaceae). Int J Mol Sci. 2023;24:10034.
69. Li X, Ding Z, Miao H, Bao J, Tian X. Complete chloroplast genome studies of different apple varieties indicated the origin of modern cultivated apples from *Malus sieversii* and *Malus sylvestris*. PeerJ. 2022;10:e13107.
70. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull. 1987;19:11–5.
71. Jin JJ, Yu WB, Yang JB, Song Y, de Pamphilis CW, Yi TS, et al. GetOrganelle: a fast and versatile toolkit for accurate de Novo assembly of organelle genomes. Genome Biol. 2020;21:1–31.
72. Krumsiek J, Arnold R, Rattei T. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinformatics. 2007;23(8):1026–8.
73. Liu S, Ni Y, Li J, Zhang X, Yang H, Chen H, et al. CPGview: a package for visualizing detailed chloroplast genome structures. Mol Ecol Resour. 2023;23:694–704.
74. Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, et al. GeSeq – versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 2017;45:W6–11.
75. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28:1647–9.
76. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. Vista: computational tools for comparative genomics. Nucleic Acids Res. 2004;32:W273-9.
77. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, et al. PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Mol Ecol Resour. 2020;20(1):348–55.
78. Xiang CY, Gao FL, Jakovlić I, Lei HP, Hu Y, Zhang H, et al. Using PhyloSuite for molecular phylogeny and tree-based analyses. iMeta. 2023;2:e87.
79. Rozas J, Ferrer-Mata A, Sanchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol. 2017;34:3299–302.
80. Kurtz S. Schleiermacher. REPuter: fast computation of maximal repeats in complete genomes. Bioinf (Oxford England). 1999;15:426–7.
81. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. Bioinformatics. 2017;33:2583–5.
82. Kolde R, _pheatmap. Pretty Heatmaps_. R package version 1.0.12, 2019. https://CRAN.R-project.org/package=pheatmap.
83. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3.
84. Nguyen LT, Schmidt HA, Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32:268–74.
85. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61:539–42.

Publisher's Note